

Asymptotic optimality conditions for linear semi-infinite programmingY. Liu^{a*} and M.A. Goberna^b^a*School of Mathematics and Geospatial Sciences, RMIT University, Melbourne, Australia;* ^b*Department of Statistics and Operations Research, University of Alicante, Alicante, Spain*

(September 22, 2014)

In this paper, the classical KKT, complementarity, and Lagrangian saddle point conditions are generalized to obtain equivalent conditions characterizing the optimality of a feasible solution to a general linear semi-infinite programming problem without constraint qualifications. The method of this paper differs from the usual convex analysis methods and its main idea is rooted in some fundamental properties of linear programming.

Keywords: Linear semi-infinite programming; inclusive cone; inclusive region; optimality condition

AMS Subject Classification: MSC 90C34; MSC 90C46; MSC 90C22; MSC 90C25

1. Introduction

Semi-infinite programming (SIP) has been an attractive area of research in optimization for decades. It has important applications in industry, economics, and science and engineering (see [13] for details). Particularly, it has interesting applications in other classes of optimization problems such as semi-definite programming and optimal control ([9, 16, 18]). At the same time, convex programming and convex semi-infinite programming problems can be formulated as linear semi-infinite programming (LSIP) problems through linearization. It is reasonable to expect that ideas and methods for LSIP also provide insight into other classes of optimization problems.

The optimality theory of LSIP has been developed for different classes of LSIP problems characterized by various constraint qualifications (CQs). Extensive summary of the most common and most important classes of LSIP problems is available from [2, 4]. Of all these classes of LSIP problems, the largest class seems to be the one specified by the so-called locally Farkas-Minkowski (LFM) CQ (note: LFM can be further relaxed by narrowing the locality down to an interested point [15]). It is well known that the classical KKT, complementarity, and Lagrangian saddle point conditions are equivalent and that they each characterize the optimality for LFM LSIP. Other classes of LSIP problems are also discussed in [2–4] and [8]. An asymptotic optimality condition was developed in [1] for linear infinite programming without any CQs. For other types of semi-infinite programming problems,

*Corresponding author. Email: yanqun.liu@rmit.edu.au

optimality conditions are discussed in [6, 13, 14] and [17], as well as relevant works cited therein.

The fundamental theorem of linear programming (LP) was recently extended to continuous LSIP in [12]. In brief, this new result establishes that under the Slater condition, a continuous LSIP problem has an optimal solution if and only if it has one at a generalized corner point of the feasible region. In addition, based on this result, an efficient numerical method was developed for continuous LSIP. Motivated by this new development, we explored the possibility of extending the method of [12] to other classes of problems. This paper presents optimality results for general LSIP problems without any CQs. The method of this paper is one of some geometrical nature and the results have clear geometric interpretations.

Consider the following general LSIP problem.

$$\begin{aligned} \min \quad & c^T x \\ \text{s. t.} \quad & a(t)x \leq b(t), \quad t \in T, \end{aligned}$$

where $x \in \mathbb{R}^n$ (the space of $n \times 1$ vectors) is the decision variable, $c \in \mathbb{R}^n$ is a given objective vector, T is an arbitrarily given infinite index set that is not required to have any topological structure, and $a : T \rightarrow \mathbb{R}_n$ (the space of $1 \times n$ vectors) and $b : T \rightarrow \mathbb{R}$ are real functions on T .

The above problem is completely determined by its objective vector c , its constraint functions a and b and their common domain T . In this paper, the above problem will be referred to as problem $P(c; a, b, T)$.

Consider the constraint vector function a in problem $P(c; a, b, T)$. If $a(t_0) = 0$ for some $t_0 \in T$, the corresponding constraint either will never be satisfied (when $b(t_0) < 0$), causing an infeasible problem, or will always be satisfied, making itself valueless. This claim is valid as far T doesn't have any topological structure or its topology isn't a concern, which is exactly the case for this paper, though in general removing t_0 from T may destroy some of its topological properties such as compactness. Thus, we rule out the extreme case in which $a(t)$ vanishes at some $t \in T$ and suppose throughout this paper, without loss of generality, that the constraint normal vectors $a(t)$ for all $t \in T$ are unit vectors. For convenience, we also suppose that the objective vector c is a unit vector unless mentioned otherwise. The feasible region and the optimal solution set of problem $P(c; a, b, T)$ are denoted by \mathcal{F} and \mathcal{F}^* , respectively.

For convenience, the notation $P(c; a, b, T)$ for an LSIP problem is frequently used to denote LP problems by replacing T with a finite index set Ω . In this paper, some LP problems may appear in the following conventional form

$$\begin{aligned} \min \quad & c^T x \\ \text{s. t.} \quad & Ax \leq d, \end{aligned}$$

where $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, and $d \in \mathbb{R}^m$ are given. This problem will be referred to as problem $LP(c; A, d)$.

We write a matrix A as $A = [a^1; a^2; \dots; a^m]$ to indicate that A has rows a^1, a^2, \dots, a^m , and a matrix B as $B = [b^1, b^2, \dots, b^n]$ to indicate that B has columns b^1, b^2, \dots, b^n .

The outline of this paper is as follows. Following this introductory section, the concepts of inclusive cone and inclusive region are introduced in Section 2, together with some known linear programming results presented in terms of inclusive cones and inclusive regions. Both concepts and results of Section 2 are fundamental to

the developments of the new results of later sections. Section 3 consists of an introduction of the so-called intersection inclusive region and its properties that are crucially important to the proof of the main results of this paper, which are given in Section 4. The first main result is Theorem 4.4 which extends the classical complementarity condition to an asymptotic form that characterizes the optimality of a feasible point without CQ. As consequences of Theorem 4.4, the classical KKT and Lagrangian saddle point conditions are also generalized to asymptotic forms. Their equivalence to the optimality of a feasible solution is stated in Theorem 4.5. Two illustrative examples are given in this section to show how the new results work when the classical optimality conditions do not apply. Finally, Section 5 contains a brief summary and some comments.

2. Inclusive Cone

The underlying idea of the main results of this paper depends on the concepts of inclusive cone and inclusive region (also called ladder) which were initially introduced in [10] for LP and in [12] for LSIP. For the purpose of this paper, they are slightly extended here so that these concepts are not necessarily related to a specific LP or LSIP problem. Some related properties to be used later are included in this section.

For given vectors $a^j \in R^n$ (or R_n), $j = 1, 2, \dots, k$, we denote by

$$\text{cone}\{a^1; a^2; \dots; a^k\}$$

the smallest convex cone containing $\{a^1; a^2; \dots; a^k\}$.

Definition 1 Let c and $d = [d_1; d_2; \dots; d_n]$ be given vectors in \mathbb{R}^n , and

$$A = [a^1; a^2; \dots; a^n], \quad a^j \in \mathbb{R}_n, \quad j = 1, 2, \dots, n,$$

be a given $n \times n$ invertible matrix. If

$$-c^T \in \text{cone}\{a^1, a^2, \dots, a^n\},$$

$\text{cone}\{a^1, a^2, \dots, a^n\}$ is called an *inclusive cone* of $-c$, and the set $\{x \mid Ax \leq d\}$ is called an *inclusive region* (or a *ladder*) associated with c , denoted by

$$L(c; A, d) = \{x \mid Ax \leq d\}.$$

The unique solution of the linear system $Ax = d$ is the unique *vertex* of $L(c; A, d)$.

For a given inclusive region, its vertex is its only corner point. The following lemma is a direct consequence of Theorem 2.5(e) in [10], which is used in the proof of Lemma 3.1.

LEMMA 2.1 *Let $L(c; A, d)$ be an inclusive region associated with c and x^0 be the vertex of $L(c; A, d)$, and $\alpha x \leq \beta$ is a linear inequality with $\alpha \in R_n$ and $\beta \in R$. If $\alpha x^0 > \beta$ and the inequalities $[A; \alpha]x \leq [d; \beta]$ is consistent, then there is some j_0 , $1 \leq j_0 \leq n$, such that $L(c; \bar{A}, \bar{d})$ is an inclusive region associated with c , where*

$$\bar{A} = [a^1; a^2; \dots; a^{j_0-1}; \alpha; a^{j_0+1}; \dots; a^n],$$

and

$$\bar{d} = [d_1; d_2; \cdots; b_{j_0-1}; \beta; d_{j_0+1}; \cdots; d_n].$$

□

Definition 2 ([10, 12]) Consider a problem $P(c; a, b, \Omega)$ where Ω is either a finite or an infinite set. The convex cone generated by n linearly independent constraint normal vectors $a(t_i)^T$, $t_i \in T$, $i = 1, 2, \cdots, n$, is called an *inclusive cone for problem* $P(c; a, b, \Omega)$ if

$$-c \in \text{cone} \{a(t_1)^T, a(t_2)^T, \cdots, a(t_n)^T\}.$$

In this case, the corresponding set of indices $I = \{t_1, t_2, \cdots, t_n\}$ is called the *generator of the inclusive cone* which is then denoted by $N(I)$.

If $N(I)$ is an inclusive cone for problem $P(c; a, b, \Omega)$, then according to Definition 1, for

$$A_I = [a(t_i); a(t_2); \cdots; a(t_n)]$$

and

$$b_I = [b(t_i); b(t_2); \cdots; b(t_n)],$$

$L(c; A_I, b_I)$ is an inclusive region associated with c . This inclusive region will often be denoted by $L(c; a, b, I)$ or simply by $L(I)$.

The following theorem, useful in our later development, is from [10]. It is parallel to the fundamental theorem of LP and bears both theoretical and computational importance ([10–12]).

THEOREM 2.2 ([10]) Consider a linear programming problem $P(c; a, b, \Omega)$ in \mathbb{R}^n . Suppose that the set of constraint normal vectors $\{a(t) : t \in \Omega\}$ contains a basis for \mathbb{R}_n . Then, the following statements are true.

- (a) Problem $P(c; a, b, \Omega)$ has an optimal solution if and only if it has one at the vertex of some inclusive region of problem $P(c; a, b, \Omega)$.
- (b) Problem $P(c; a, b, \Omega)$ has no optimal solution if and only if it is either infeasible or has no inclusive region (or, inclusive cone). □

From the classical KKT theorem, it is easy to see that the following is also true.

LEMMA 2.3 Let x^* be a corner point of the feasible region of a linear programming problem $P(c; a, b, \Omega)$. Then, x^* is optimal if and only if x^* is the vertex of an inclusive region of problem $P(c; a, b, \Omega)$. □

3. Intersection Inclusive Cone

In this section, we discuss the so called intersection inclusive cone and intersection inclusive region obtained by restricting a given inclusive region to a special hyperplane. Some properties of the intersection inclusive region are given here and will be used in the next section.

Throughout this section, we suppose that $L(c; A, d)$ is an inclusive region associated with c , and x^0 is its vertex, where

$$A = [a^1; a^2; \cdots; a^n], \quad a^j \in \mathbb{R}_n, \quad j = 1, 2, \cdots, n,$$

and

$$d = [d_1; d_2; \cdots; d_n], \quad d_j \in \mathbb{R}, \quad 1 \leq j \leq n.$$

We see that an inclusive region $L(c; A, d)$ is a conical region having n edges. The j -th edge of $L(c; A, d)$, denoted by E_j , is given by

$$E_j = \{x \in \mathbb{R}^n \mid a^j x \leq d_j \text{ and } a^s x = d_s \text{ for all } 1 \leq s \leq n, s \neq j\}.$$

The unit vector in the recession direction of $L(c; A, d)$ that is parallel to E_j is called the *direction vector* of E_j and is denoted by e^j .

For $1 \leq l \leq n$, let

$$\tilde{A} = [a^1; a^2; \cdots; a^{l-1}; -e^l; a^{l+1}; a^{l+2}; \cdots; a^n], \quad (1)$$

$$A(l) = [a^1; a^2; \cdots; a^{l-1}; a^{l+1}; \cdots; a^n], \quad (2)$$

$$\tilde{d} = [d_1; d_2; \cdots; d_{l-1}; -e^l x^*; d_{l+1}; d_{l+2}; \cdots; d_n], \quad (3)$$

and

$$d(l) = [d_1; d_2; \cdots; d_{l-1}; d_{l+1}; \cdots; d_n]. \quad (4)$$

If for some $1 \leq j \leq n$, e^j satisfies

$$c^T e^j = \min_{1 \leq s \leq n} c^T e^s,$$

we call the edge E_j a *minimal edge* of $L(c; A, d)$. We note that a minimal edge is not unique in general.

Suppose that E_l is a minimal edge of $L(c; A, d)$ with direction vector e^l . For any $x^* \in L(c; A, d)$, let

$$H(x^*, l) = \{x \in \mathbb{R}^n \mid e^l x = e^l x^*\}$$

be the hyperplane passing through x^* and normal to e^l . Clearly, the line extending E_l in \mathbb{R}^n is cut by $H(x^*, l)$ at

$$\bar{x}^0 = x^0 + ((x^* - x^0)^T e^l) e^l. \quad (5)$$

If $x^* \neq \bar{x}^0$, the unit vector v^1 defined by

$$v^1 = (\bar{x}^0 - x^*) / \|\bar{x}^0 - x^*\| \quad (6)$$

is normal to e^l , and hence can be extended into an orthonormal basis

$$\{v^1, v^2, \cdots, v^{n-1}\}$$

for the subspace

$$H(x^*, l) - x^* = \{e^l\}^\perp \subset \mathbb{R}^n.$$

For any $x \in H(x^*, l)$, since $x - x^* \in \{e^l\}^\perp$, there exists a unique

$$y = [y_1; y_2; \dots; y_{n-1}] \in \mathbb{R}^{n-1}$$

such that

$$x - x^* = y_1 v^1 + y_2 v^2 + \dots + y_{n-1} v^{n-1} = Vy,$$

where $V = [v^1, v^2, \dots, v^{n-1}]$. Thus,

$$x = Vy + x^*. \quad (7)$$

Note that V is an orthogonal matrix. If V is considered a linear transformation from \mathbb{R}^{n-1} to \mathbb{R}^n , however, V^T is only a left inverse of V . In addition, we see that (7) defines an invertible mapping from \mathbb{R}^{n-1} to $H(x^*, l)$.

Some useful properties are given in the following two lemmas.

LEMMA 3.1 *Let E_l be a minimal edge of $L(c; A, d)$ with direction vector e^l . Then $L(c; \bar{A}, \bar{d})$, where \bar{A} and \bar{d} are defined by (1) and (3), is an inclusive region associated with c with vertex \bar{x}^0 given by (5).*

Proof. We need only to show that the columns of \bar{A}^T generate an inclusive cone for c . To this end, we consider the LP problem $\text{LP}(c; \bar{A}, \bar{d})$, where

$$\bar{A} = [A; -(e^l)^T] \quad \text{and} \quad \bar{d} = [b; -(e^l)^T x^0 - 1].$$

Note that x^0 violates the constraint

$$-(e^l)^T x \leq -(e^l)^T x^0 - 1.$$

Thus, according to Lemma 2.1, a new ladder is available from $L(c; A, d)$ by replacing one of the rows of A with $-(e^l)^T$ and replacing the corresponding component of d with $-(e^l)^T x^0 - 1$. Clearly, the vertex of the new ladder, denoted by x^* , is feasible to problem $\text{LP}(c; \bar{A}, \bar{d})$. Theorem 2.2 indicates that x^* , as a feasible ladder vertex, is an optimal solution of problem $\text{LP}(c; \bar{A}, \bar{d})$. Clearly, x^* is the intersection of the hyperplane

$$-(e^l)^T x = -(e^l)^T x^0 - 1 \quad (8)$$

with one of the edges of $L(c; A, d)$.

Suppose that x^* is the intersection of the hyperplane (8) with E_j . It is obvious that $x^* = x^0 + \sigma e^j$ for some $\sigma \geq 1$. Since $c^T e^j \geq 0$ and E_l is a minimal edge, we see that

$$c^T(x^0 + e^l) \leq c^T(x^0 + e^j) \leq c^T x^*.$$

This implies that the intersection of the hyperplane (8) with E_l , $x^0 + e^l$, is an optimal solution of problem $LP(c; \bar{A}, \bar{d})$. Therefore, according to Lemma 2.3, $L(c; \tilde{A}, \hat{d})$ is an inclusive region, where

$$\hat{d} = [d_1; d_2; \dots; d_{l-1}; -e^l x^* - 1; d_{l+1}; d_{l+2}; \dots; d_n].$$

Thus, the columns of \tilde{A}^T generates an inclusive cone associated with c . \square

LEMMA 3.2 *Suppose that the following conditions are satisfied:*

- (a) $x^* \in L(c; A, b) \setminus E_l$,
- (b) E_l is a minimal edge of $L(c; A, d)$,
- (c) \bar{x}^0 is the intersection of $H(x^*, l)$ with the line extending E_l , and
- (d) $V = [v^1, v^2, \dots, v^{n-1}]$, where $\{v^1, v^2, \dots, v^{n-1}\}$ is an orthonormal basis of $H(x^*, l) - x^*$, of which v^1 is given by (6).

Then, $L(\tau; \Lambda, \beta)$ is an inclusive region associated with τ in \mathbb{R}^{n-1} , where $\tau = V^T c$, $\Lambda = A(l)V$, and $\beta = d(l) - A(l)x^*$. Furthermore, $L(\tau; \Lambda, \beta)$ contains the zero vector of \mathbb{R}^{n-1} and has vertex $y^0 = [\|x^0 - x^*\|; 0; \dots; 0]$. \square

Proof. From Lemma 3.1, $L(c; \tilde{A}, \tilde{d})$ is an inclusive region associated with c . Thus, there are $\lambda_j \geq 0$, $j = 1, 2, \dots, n$, such that

$$-c = \sum_{j \neq l, j=1}^n \lambda_j (a^j)^T + \lambda_l (-e^l). \quad (9)$$

Let $\bar{c} = c - (c^T e^l) e^l$ be the projection of c onto $H(x^*, l) - x^*$ ($= \{e^l\}^\perp$). We know that there exists $\tau \in \mathbb{R}^{n-1}$ such that $\bar{c} = V\tau$, which implies that

$$\tau = V^T \bar{c} = V^T (c - (c^T e^l) e^l) = V^T c \quad (10)$$

since

$$V^T e^l = [v^1, v^2, \dots, v^{n-1}]^T e^l = 0 \in \mathbb{R}^{n-1}.$$

Now, when x is restricted to the hyperplane $H(x^*, l)$, according to (7), the constraints $a^j x \leq b^j$ for $1 \leq j \leq n$, $j \neq l$, become

$$a^j (Vy + x^*) \leq d_j, \quad 1 \leq j \leq n, \quad j \neq l,$$

or

$$a^j Vy \leq d_j - a^j x^*, \quad 1 \leq j \leq n, \quad j \neq l,$$

which are constraints in \mathbb{R}^{n-1} . The set

$$\{(a^j V)^T \mid 1 \leq j \leq n, \quad j \neq l\} \subset \mathbb{R}^{n-1}$$

is obviously linearly independent as its image under V is the basis

$$\{(a^j)^T \mid 1 \leq j \leq n, \quad j \neq l\}$$

for $H(x^*, l) - x^*$. In addition, it follows from (9) and (10) that

$$\begin{aligned}
 -\tau &= -V^T c \\
 &= V^T \left(\sum_{j \neq l, j=1}^n \lambda_j (a^j)^T + \lambda_l (-e^l) \right) \\
 &= \sum_{j \neq l, j=1}^n \lambda_j (a^j V)^T.
 \end{aligned} \tag{11}$$

Equation (11) and the linear independence of $\{a^j V \mid 1 \leq j \leq n, j \neq l\}$ imply that $L(\tau; \Lambda, \beta)$ is an inclusive region associated with τ in \mathbb{R}^{n-1} , where $\tau = V^T c$, $\Lambda = A(l)V$, and $\beta = d(l) - A(l)x^*$.

Now, the vertex of $L(\tau; \Lambda, \beta)$ is the unique solution y^0 of the system $\Lambda y = \beta$, which means that

$$A(l)(Vy^0 + x^*) = d(l).$$

Hence, the point $Vy^0 + x^*$, also satisfying

$$e^l(Vy^0 + x^*) = e^l x^*,$$

is a solution to $\tilde{A}x = \tilde{d}$. Thus, $Vy^0 + x^* = \bar{x}^0$, which together with the fact that

$$(v^j)^T(\bar{x}^0 - x^*) = 0, \quad 1 \leq j \leq n-1,$$

implies that

$$\begin{aligned}
 y^0 &= V^T(\bar{x}^0 - x^*) \\
 &= [(\bar{x}^0 - x^*)/\|\bar{x}^0 - x^*\|, v^2, \dots, v^{n-1}]^T(\bar{x}^0 - x^*) \\
 &= [\|\bar{x}^0 - x^*\|; 0; \dots; 0]
 \end{aligned} \tag{12}$$

Finally, it is true that $0 \in L(\tau; \Lambda, \beta) \subset \mathbb{R}^{n-1}$ as

$$\Lambda 0 = A(l)V0 = A(l)0 = A(l)(x^* - x^*) \leq d(l) - A(l)x^* = \beta.$$

The proof is complete. \square

Definition 3 If $x^* \in L(c; A, d)$ and the l -th edge of $L(c; A, d)$ is a minimal edge, we call the $n-1$ dimensional inclusive region $L(\tau; \Lambda, \beta)$ given by Lemma 3.2 the *intersection inclusive region* of $L(c; A, d)$ by $H(x^*, l)$, where $H(x^*, l)$ is the hyperplane passing through x^* and normal to e^l . The inclusive cone associated with $L(\tau; \Lambda, \beta)$ is said to be the corresponding *intersection inclusive cone*.

We note that the intersection inclusive region $L(\tau; \Lambda, \beta)$ depends on the choice of the matrix V . In addition, if the edge E_l is not minimal, the quantities τ , Λ and β can still be defined but Λ and β do not determine an inclusive region associated with τ .

4. Optimality

In this section, equivalent optimality conditions are developed for LSIP without assuming any CQs. Several existing important optimality theorems are discussed first, together with two illustrative examples. This is followed by the new optimality results and their proof. As an application, an alternative proof of the asymptotic optimality theorem for LSIP [1] is derived from the main results of this paper.

Problem $P(c; a, b, T)$ is said to be locally Farkas-Minkowski (LFM in brief) at $x^* \in \mathcal{F}$ if every linear consequence of $\{a(t)x \leq b(t), t \in T\}$ binding at x^* is a linear consequence of a finite subsystem. It is said to be Slater if there exists a Slater point. It is known that for a Slater problem $P(c; a, b, T)$ if T is a compact topological space and a and b are continuous on T then $P(c; a, b, T)$ is LFM. Both LFM and Slater LSIP problems are important in the classical optimality theory for LSIP, which depend on the so-called generalized sequences over T , that is, functions from T to \mathbb{R} . A generalized sequence $\lambda : T \rightarrow \mathbb{R}$ is said to be a generalized finite sequence on T if its support given by $\text{supp}(\lambda) = \{t \in T \mid \lambda(t) \neq 0\}$ is a finite set. Let $\mathbb{R}^{(T)}$ denote the linear space of all generalized finite sequences on T , and $\mathbb{R}_+^{(T)}$ the positive cone of $\mathbb{R}^{(T)}$. For given $f : T \rightarrow \mathbb{R}^m$ and $\lambda \in \mathbb{R}^{(T)}$ with

$$\text{supp}(\lambda) = \{t_1, t_2, \dots, t_p\},$$

we define

$$\sum_{t \in T} \lambda(t)f(t) = \begin{cases} \sum_{i=1,2,\dots,p} \lambda(t_i)f(t_i), & \text{if } p \geq 1, \\ 0, & \text{if } p = 0 \text{ (i.e., } \text{supp}(\lambda) = \emptyset \text{).} \end{cases}$$

Following [15], the optimality theorem under the LFM CQ can be stated as follows:

THEOREM 4.1 (Optimality theorem with CQ) *Assume that the constraint system of problem $P(c; a, b, T)$ is LFM at $x^* \in \mathcal{F}$. Then the following statements are equivalent:*

- (i) $x^* \in \mathcal{F}^*$.
- (ii) (KKT condition) $-c \in \mathcal{A}(x^*)$, where

$$\mathcal{A}(x^*) = \text{cone} \{a(t)^T \mid t \in T \text{ and } a(t)x^* = b(t)\}$$

is the active cone at x^ .*

- (iii) (complementarity condition) *There exist $t_j \in T$ and $\lambda_j \geq 0$, $j = 1, 2, \dots, k$, such that*

$$-c = \sum_{j=1}^k \lambda_j a(t_j)^T$$

and

$$\lambda_j(a(t_j)x^* - b(t_j)) = 0, \quad j = 1, 2, \dots, k;$$

(iv) (Lagrangian saddle point condition) There exists $\lambda^* \in \mathbb{R}_+^{(T)}$ such that

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*), \quad \text{for all } x \in \mathbb{R}^n \text{ and all } \lambda \in \mathbb{R}_+^{(T)}, \quad (13)$$

where the Lagrangian $L(x, \lambda)$ is given by

$$L(x, \lambda) = c^T x + \sum_{t \in T} \lambda(t)(a(t)x - b(t)).$$

All known optimality conditions without any CQ are in asymptotic form. The next optimality theorem characterizing the optimality of a feasible solution without any CQ comes from [1] (Corollary 5).

THEOREM 4.2 (Optimality theorem without CQ) *A point $x^* \in \mathcal{F}$ is optimal if and only if there exists a sequence $\{(\lambda_i, \varepsilon_i)\} \subset \mathbb{R}_+^{(T)} \times \mathbb{R}$ such that*

$$\sum_{t \in T} \lambda_i(t)b(t) \leq \varepsilon_i - c^T x^*, \quad \text{for } i = 1, 2, \dots, \quad (14)$$

and

$$\left(\sum_{t \in T} \lambda_i(t)a(t), \varepsilon_i\right) \rightarrow (-c, 0_+). \quad (15)$$

It's also worth mentioning the result of [5] (Corollary 5.5) which provides a sufficient condition of the optimality of a feasible x^* for LSIP without CQ. That result is given in terms of the so-called cone of extended active constraints at x^* .

The following are two examples for which the conditions (ii)–(iv) of Theorem 4.1 fail to be necessary conditions for (i). These examples will also be used later in this section to illustrate the new results of this paper.

EXAMPLE 1 $c = [1; 0]$, $a(t) = [-\sin(t), \cos(t)]$, $b(t) = \sin^2(t)$, $T = \{t \mid 0 \leq t \leq \pi\}$.

It is easy to see that the feasible region of this problem is $\mathcal{F} = \{[x_1; 0] \mid x_1 \geq 0\}$. The unique optimal solution is $x^* = [0; 0]$.

The inequality $-x_1 \leq 0$, which gives a supporting hyperplane $x_1 = 0$ of \mathcal{F} , is a consequence of the given constraint system, but not a consequence of any finite constraint subsystem. This means that the LFM CQ is not satisfied at x^* .

At the optimal solution $x^* = [0; 0]$, the only active constraints are those two that correspond to $t = 0$ and $t = \pi$, respectively. The classical KKT, complementarity, and the Lagrange saddle point conditions in Theorem 4.1 are not satisfied. In fact, it is obvious that the classical KKT and the complementarity conditions fail to hold at x^* . To see that the saddle point condition also fails, we note that the first inequality in (13) leads to conclusion that the support of λ^* must be a subset of $\{0, \pi\}$. However, for any of the three possible cases for $\text{supp}(\lambda^*)$, namely $\{0\}$, $\{\pi\}$, and $\{0, \pi\}$, one can check that the second inequality is violated at $x = [-1; 0]$.

EXAMPLE 2 $c = [0; 1]$, $a(t) = [\sin(t), -\cos(t)]$, $b(t) = 0$, $T = \{t \mid 0 < t < \pi/2\}$.

For this example, the feasible region is $\mathcal{F} = \{[x_1; x_2] \mid x_1 \leq 0, x_2 \geq 0\}$. There are multiple solutions and the optimal solution set is $\mathcal{F}^* = \{[x_1; 0] \mid x_1 \leq 0\}$.

The inequality $-x_2 \leq 0$, which gives a supporting hyperplane $x_2 = 0$ of \mathcal{F} , is a consequence of the given constraint system, but it is not a consequence of any

finite constraint subsystem. The LFM CQ is thus not satisfied.

At the optimal solution $[0; 0]$, all constraints are active. At any other optimal solution $[x_1^*; 0]$ ($x_1^* < 0$), there is simply no active constraint. The classical KKT, complementarity, and the Lagrange saddle point conditions are not satisfied at any optimal solution.

The above examples show that the LFM CQ cannot be completely removed for conditions (ii)–(iv) in Theorem 4.1 to be necessary for optimality. The remaining of this section is devoted to obtain generalized versions of the classical KKT, complementarity, and Lagrangian saddle point conditions that characterize the optimality of a feasible solution for LSIP without any CQs.

In the following, we consider a sequence of inclusive regions

$$L(c^i; A_i, b^i), \quad i = 1, 2, \dots, \quad (16)$$

with corresponding sequence of vertices x^i , $i = 1, 2, \dots$, where

$$A_i = [a^{i,1}; a^{i,2}; \dots; a^{i,n}], \quad a^{i,j} \in \mathbb{R}_n, \quad 1 \leq j \leq n, \quad (17)$$

and

$$b^i = [b_1^i; b_2^i; \dots; b_n^i], \quad b_j^i \in \mathbb{R} \quad j = 1, 2, \dots, n. \quad (18)$$

We assume that the following properties are satisfied.

$$\|c^i\| = \|a_j^i\| = 1, \quad i = 1, 2, \dots, \quad 1 \leq j \leq n, \quad (19)$$

$$c^i \rightarrow c^0 \quad (i \rightarrow \infty), \quad \text{for some } c^0 \in \mathbb{R}^n, \quad (20)$$

$$x^i \rightarrow x^0 \quad (i \rightarrow \infty), \quad \text{for some } x^0 \in \mathbb{R}^n. \quad (21)$$

Sequences of inclusive regions satisfying (19)–(21) are of special importance to the new optimality results. The following is a related result that will be used in the proof of the first main result, Theorem 4.4.

LEMMA 4.3 *Consider the sequence of inclusive regions given by (16)–(18) with corresponding sequence of vertices $\{x^i\}$. If conditions (19)–(21) are satisfied and there exists $x^* \in \cap_{i=1}^\infty L(c^i; A_i, b^i)$ such that*

$$(c^0)^T(x^0 - x^*) = 0, \quad (22)$$

$$d = \|x^0 - x^*\| > 0. \quad (23)$$

Then, there exist an infinite sequence of integers, denoted by I , and corresponding sequences $\{\tilde{c}^i\}_{i \in I}$ and $\{\tilde{x}^i\}_{i \in I}$ in \mathbb{R}^n , and indices $1 \leq s(i, j) \leq n$, for $i \in I$ and $j = 1, 2, \dots, k$ ($1 \leq k \leq n$) such that

$$\tilde{c}^i - c^i \rightarrow 0 \quad (i \in I, \quad i \rightarrow \infty), \quad (24)$$

$$\tilde{x}^i \rightarrow x^* \quad (i \in I, \quad i \rightarrow \infty), \quad (25)$$

$$a^{i,s(i,j)} \tilde{x}^i = b_{s(i,j)}^i, \quad i \in I, \quad j = 1, 2, \dots, k, \quad (26)$$

and for each $i \in I$ there exist $\lambda_j^i \geq 0$, $j = 1, 2, \dots, k$ such that

$$-\tilde{c}^i = \sum_{j=1}^k \lambda_j^i (a^{i,s(i,j)})^T, \quad (27)$$

where $a^{i,s(i,j)}$, $j = 1, 2, \dots, k$, are linearly independent.

Proof. When $n = 1$, (19) and (20) mean $|c^0| = 1$. It follows that conditions (22) and (23) cannot be satisfied at the same time. Thus, the lemma is true for $n = 1$.

To prove the lemma for $n \geq 2$, we use induction on the dimension n .

When $n = 2$, we have $A_i = [a^{i,1}; a^{i,2}]$ and $b^i = [b_1^i; b_2^i]$. As x^i minimizes $(c^i)^T x$ over $L(c^i; A_i, b^i)$, we see that $(c^i)^T x^* \geq (c^i)^T x^i$. Since

$$(c^i)^T (x^* - tc^i) \rightarrow -\infty \quad (t \rightarrow +\infty),$$

it follows that

$$x^* - tc^i \notin L(c^i; A_i, b^i) \quad \text{for sufficiently large } t.$$

Thus, there exist index $s(i, 1) \in \{1, 2\}$ and constants α_i and ε_i with $\alpha_i \geq \varepsilon_i \geq 0$ such that $\hat{x}^i = x^* - \alpha_i c^i$ and $\tilde{x}^i = x^* - \varepsilon_i c^i$ satisfy

$$a^{i,s(i,1)} \tilde{x}^i = b_{s(i,1)}^i, \quad (28)$$

$$a^{i,s(i,1)} (x^* - \varepsilon c^i) > b_{s(i,1)}^i, \quad \text{for } \varepsilon > \varepsilon_i,$$

and

$$(c^i)^T x^* - \alpha_i \|c^i\|^2 = (c^i)^T \hat{x}^i = (c^i)^T x^i. \quad (29)$$

From equation (29) and the assumptions (19)-(23), we get

$$\alpha_i = \alpha_i \|c^i\|^2 = (c^i)^T (x^* - x^i) \rightarrow 0.$$

This implies

$$0 \leq \varepsilon_i \leq \alpha_i \rightarrow 0.$$

Thus, from the definition of \hat{x}^i and \tilde{x}^i ,

$$\lim_{i \rightarrow \infty} \hat{x}^i = \lim_{i \rightarrow \infty} \tilde{x}^i = x^*. \quad (30)$$

Equation (30), together with (21), implies

$$\lim_{i \rightarrow \infty} \left[\frac{\tilde{x}^i - x^i}{\|\tilde{x}^i - x^i\|} - \frac{\hat{x}^i - x^i}{\|\hat{x}^i - x^i\|} \right] = 0. \quad (31)$$

Since $(\hat{x} - x^i) \perp (x^* - \hat{x})$, \tilde{x} is on the line segment connecting x^* and \hat{x} (see Figure 1,

where a^i represents $a^{i,s(i,1)}$, and $\|c^i\| = \|a^{i,s(i,1)}\| = 1$, we see that

$$\left\| (a^{i,s(i,1)})^T + c^i \right\| = \left\| \frac{\tilde{x}^i - x^i}{\|\tilde{x}^i - x^i\|} - \frac{\hat{x}^i - x^i}{\|\hat{x}^i - x^i\|} \right\|,$$

(31) indicates that

$$(a^{i,s(i,1)})^T + c^i \rightarrow 0 \quad (i \rightarrow \infty). \quad (32)$$

From (28), (30) and (32), the lemma holds true for $n = 2$ if we choose $I = \{1, 2, \dots\}$ and $\tilde{c}^i = -(a^{i,s(i,1)})^T$ for all $i \in I$.

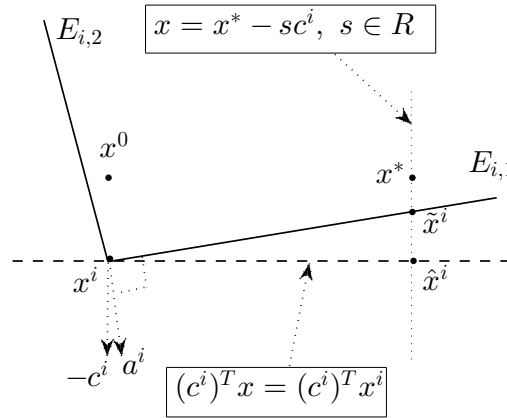


Figure 1. Explanation of the proof of Lemma 4.3, case $n = 2$.

Suppose that the lemma is true for $n = p \geq 2$. Consider $n = p + 1$. For each integer i , we assume without loss of generality that the n -th edge $E_{i,n}$ of $L(c^i; A_i, b^i)$ is one of its minimal edges, and the direction vector of $E_{i,n}$ is simply denoted by e^i . The hyperplane passing through x^* and normal to e^i is denoted by $H_i(x^*, n)$.

Let

$$\tilde{A}_i = [a^{i,1}; a^{i,2}; \dots; a^{i,p}; -e^i], \quad i = 1, 2, \dots,$$

and

$$\tilde{b}^i = [b_1^i; b_2^i; \cdots; b_p^i; -e^i x^*], \quad i = 1, 2, \cdots.$$

Since $E_{i,n}$ is a minimal edge of $L(c^i; A_i, b^i)$, we see that, for each integer i , problem $P(c^i; \tilde{A}_i, \tilde{b}^i)$ has optimal solution at its corner point. Thus, $L(c^i; \tilde{A}_i, \tilde{b}^i)$ is an inclusive region associated with c^i . The vertex of $L(c^i; \tilde{A}_i, \tilde{b}^i)$ is denoted by \bar{x}^i .

To complete the proof, we consider two complementary situations.

Case I. There exists an infinite sequence of integers, denoted by I , such that $\bar{x}^i = x^*$ for all $i \in I$. In this case, for each $i \in I$, we choose $\tilde{x}^i = \bar{x}^i$ and $\tilde{c}^i = c^i$. Then, it can be seen that the lemma holds true according to Lemma 3.1.

Case II. There are at most finitely many i 's such that $\bar{x}^i = x^*$. Then, $\bar{x}^i \neq x^*$ if i is sufficiently large.

We define $v^{i,1}$ by

$$v^{i,1} = (\bar{x}^i - x^*) / \|\bar{x}^i - x^*\|$$

and extend $\{v^{i,1}\}$ into an orthonormal basis for $H_i(x^*, n) - x^* = \{e^i\}^\perp$:

$$\{v^{i,1}, v^{i,2}, \dots, v^{i,p}\}.$$

Then we define a matrix V_i by

$$V_i = [v^{i,1}, v^{i,2}, \dots, v^{i,p}].$$

By Lemma 3.2, the intersection inclusive region $L(\tau^i; \Lambda_i, \beta^i)$ of $L(c^i; A_i, b^i)$ by $H_i(x^*, n)$ is well defined. The intersection inclusive region $L(\tau^i; \Lambda_i, \beta^i)$, and its vertex y^i satisfy, for $i = 1, 2, \dots$,

$$\tau^i = V_i^T \bar{c}^i = V_i^T (c^i - ((c^i)^T e^i) e^i) = V_i^T c^i, \quad (33)$$

$$\Lambda_i = A_i(n) V_i = [a^{i,1}; a^{i,2}; \dots; a^{i,p}] V_i, \quad (34)$$

$$\beta^i = b^i(n) - A_i(n) x^* = [b_1^i; b_2^i; \dots; b_p^i] - A_i(n) x^*, \quad (35)$$

$$y^i = V_i^T (\bar{x}^i - x^*) = [\|\bar{x}^i - x^*\|; 0; \dots; 0], \quad (36)$$

and

$$0 \in \cap_{i=1}^\infty L(\tau^i; \Lambda_i, \beta^i). \quad (37)$$

If there is a sequence of integers I such that the corresponding sequence $\{y^i\}_{i \in I}$ satisfies

$$y^i \rightarrow 0 \quad (i \in I, i \rightarrow \infty), \quad (38)$$

we choose $\tilde{x}^i = \bar{x}^i$ and $\tilde{c}^i = \bar{c}^i$ for all $i \in I$, where \bar{c}^i is the projection of c^i onto $H_i(x^*, n)$. Since, according to (36), (38) implies

$$\bar{x}^i \rightarrow x^* \quad (i \in I, i \rightarrow \infty),$$

thus we have

$$\tilde{x}^i \rightarrow x^* \quad (i \in I, i \rightarrow \infty), \quad (39)$$

From (22), we have

$$(c^i)^T (\bar{x}^i - x^i) \rightarrow (c^0)^T (x^* - x^0) = 0 \quad (i \in I, i \rightarrow \infty),$$

and hence

$$|(c^i)^T e^i| = |(c^i)^T ((x^i - \bar{x}^i) / \|x^i - \bar{x}^i\|)| \rightarrow 0 \quad (i \in I, i \rightarrow \infty).$$

Thus, \tilde{c}^i satisfies

$$\tilde{c}^i = \bar{c}^i = c^i - ((c^i)^T e^i) e^i \rightarrow c^0 \quad (i \in I, i \rightarrow \infty). \quad (40)$$

For each $i \in I$, \bar{x}^i hence $\tilde{x}^i (= \bar{x}^i)$ is on the line obtained by extending $E_{i,n}$. Therefore,

$$a^{i,j} \tilde{x}^i = b_j^i, \quad \text{for all } i \in I \text{ and } j = 1, 2, \dots, p. \quad (41)$$

Furthermore, for $i \in I$, as $L(\tau^i; \Lambda_i, \beta^i)$ is an inclusive region associated with τ^i , we have

$$-\tau^i = \sum_{j=1}^p \lambda_{i,j} (a^{i,j} V_i)^T.$$

Consequently, we have

$$\begin{aligned} -\tilde{c}^i &= -\bar{c}^i \\ &= V_i(-\tau^i) \\ &= V_i \sum_{j=1}^p \lambda_{i,j} (a^{i,j} V_i)^T \\ &= \sum_{j=1}^p \lambda_{i,j} (a^{i,j})^T. \end{aligned} \quad (42)$$

Expressions (39)–(42) show that, in case (38) is satisfied, the lemma is true for $n = p + 1$.

On the other hand, if (38) is not satisfied for any infinite integer sequence I , there must exist a sequence I of integers such that

$$y^i \rightarrow y^0 \neq 0 \quad (i \in I, i \rightarrow \infty). \quad (43)$$

Let $y^* = 0$ be the zero vector in \mathbb{R}^p . From (33)–(37), $\{L(\tau^i; \Lambda_i, \beta^i)\}_{i \in I}$ satisfies

$$y^* \in \cap_{i=1}^{\infty} L(\tau^i; \Lambda_i, \beta^i), \quad (44)$$

and

$$\|y^i - y^*\| \rightarrow \|y^0\| > 0 \quad (i \in I, i \rightarrow \infty). \quad (45)$$

Clearly, $\{\tau^i\}_I$ is bounded. It has a converging subsequence which is assumed without loss of generality to be $\{\tau^i\}_I$ itself and satisfies

$$\tau^i \rightarrow \tau^0 \quad (i \in I, i \rightarrow \infty) \quad (46)$$

for some $\tau^0 \in \mathbb{R}^p$. In addition, since

$$c^i - \bar{c}^i = ((c^i)^T e^i) e^i \rightarrow 0 \quad (47)$$

and

$$(\bar{c}^i)^T(\bar{x}^i - x^*) = (\bar{c}^i)^T(\bar{x}^i - x^i) + (\bar{c}^i)^T(x^i - x^*) \rightarrow 0 \quad (i \rightarrow \infty),$$

we have

$$\begin{aligned} (\tau^i)^T y^i &= (V_i^T \bar{c}^i)^T y^i \\ &= (\bar{c}^i)^T (V_i y^i) \\ &= (\bar{c}^i)^T (\bar{x}^i - x^*) \\ &= (\bar{c}^i)^T (\bar{x}^i - x^i) + (\bar{c}^i)^T (x^i - x^*) \\ &= (\bar{c}^i)^T (x^i - x^*) \\ &= (\bar{c}^i - c^i)^T (x^i - x^*) + (c^i)^T (x^i - x^*) \\ &\rightarrow 0 = (\tau^0)^T y^* \quad (i \in I, i \rightarrow \infty). \end{aligned} \tag{48}$$

Furthermore, as V_i is norm preserving on $H_i(x^*, n)$, equation (47) implies that

$$\begin{aligned} \|\tau^i\| &= \|V_i^T \bar{c}^i\| = \|\bar{c}^i\| = \|c^i - ((c^i)^T e^i) e^i\| \\ &\geq \|c^i\| - \|((c^i)^T e^i) e^i\| = 1 - |(c^i)^T e^i| \rightarrow 1 \quad (i \rightarrow \infty). \end{aligned} \tag{49}$$

We see from (43)–(49) that the sequence

$$\{L(\tau^i/\|\tau^i\|; \Lambda_i, \beta^i)\}_{i \in I} \subset \mathbb{R}^p$$

satisfies the conditions of the lemma. Therefore, according to the induction assumption, there exists a subsequence of I , still denoted by I , and corresponding sequences $\{\tilde{y}^i\}_{i \in I}$ and $\{\tilde{\tau}^i\}_{i \in I}$, and indices $1 \leq s(i, j) \leq n$, for $i \in I$ and $j = 1, 2, \dots, k$ ($1 \leq k \leq p$) such that

$$\tilde{\tau}^i \rightarrow \tau^0 \quad (i \in I, i \rightarrow \infty), \tag{50}$$

$$\tilde{y}^i \rightarrow y^* = 0 \quad (i \in I, i \rightarrow \infty), \tag{51}$$

$$a^{i, s(i, j)} V_i \tilde{y}^i = b_{s(i, j)}^i - a^{i, s(i, j)} x^*, \quad i \in I, j = 1, 2, \dots, k, \tag{52}$$

and for each $i \in I$ there exist $\mu_j^i \geq 0, j = 1, 2, \dots, l$ such that

$$-\tilde{\tau}^i = \sum_{j=1}^k \mu_j^i (a^{i, s(i, j)} V_i)^T. \tag{53}$$

For $i \in I$, let $\tilde{c}^i = V_i \tilde{\tau}^i$. Then, (33), (46) and (50) imply that

$$\|\tilde{c}^i - \bar{c}^i\| = \|V_i(\tilde{\tau}^i - \tau^i)\| \rightarrow 0 \quad (i \in I, i \rightarrow \infty).$$

This, together with the fact that $\bar{c}^i - c^i \rightarrow 0$ ($i \rightarrow \infty$), means that

$$\tilde{c}^i \rightarrow c^0 \quad (i \in I, i \rightarrow \infty). \tag{54}$$

At the same time, let \tilde{x}^i be defined as $\tilde{x}^i = V_i \tilde{y}^i + x^*$. Then (51) indicates that

$$\tilde{x}^i \rightarrow x^* \quad (i \in I, i \rightarrow \infty). \quad (55)$$

Now, $\tilde{\tau}^i = V_i^T \tilde{c}^i$ as $\tilde{c}^i = V_i \tilde{\tau}^i \in H_i(x^*, n)$. We see that (52) implies

$$a^{i,s(i,j)}(V_i \tilde{y}^i + x^*) = b_{s(i,j)}^i, \quad j = 1, 2, \dots, l,$$

and hence

$$a^{i,s(i,j)} \tilde{x}^i = b_{s(i,j)}^i, \quad j = 1, 2, \dots, k. \quad (56)$$

Equation (53) implies

$$-V_i^T \tilde{c}^i = V_i^T \sum_{j=1}^k \mu_j^i (a^{i,s(i,j)})^T,$$

which, since $\sum_{j=1}^k \mu_j^i \in H_i(x^*, n)$, leads to

$$-\tilde{c}^i = \sum_{j=1}^k \mu_j^i (a^{i,s(i,j)})^T. \quad (57)$$

Equations (54)–(57) show that the lemma is true for Case II. \square

THEOREM 4.4 *A feasible solution x^* of problem $P(c; a, b, T)$ is an optimal solution if and only if there exist, for each $i = 1, 2, \dots$, an integer $k = k(i)$ satisfying $1 \leq k \leq n$, a unit vector c^i and a point x^i , both in \mathbb{R}^n , and an index set*

$$T_i = \{t_1^i, t_2^i, \dots, t_k^i\} \subset T$$

such that

$$c^i \rightarrow c, \quad (i \rightarrow \infty), \quad (58)$$

$$x^i \rightarrow x^*, \quad (i \rightarrow \infty), \quad (59)$$

$$a(t_j^i)x^i = b(t_j^i), \quad j = 1, 2, \dots, k, \quad (60)$$

and

$$-c^i = \sum_{j=1}^k \lambda_j^i a(t_j^i)^T, \quad (61)$$

where $\lambda_j^i > 0$, and $a(t_j^i)$, $j = 1, 2, \dots, k$, are linearly independent.

Before giving a proof for this theorem, we note that the integer $k = k(i)$ in the theorem may be required to be independent of i . Furthermore, the multipliers λ_j^i may be relaxed from being positive to being non-negative.

Proof. We first prove the sufficiency. Let conditions (58)–(61) be satisfied and $\bar{x} \in \mathcal{F}$ be any given feasible solution to problem $P(c; a, b, T)$. It is clear that x^i

is an optimal solution and \bar{x} is a feasible solution to problem $P(c^i; a, b, T_i)$. Thus, $(c^i)^T x^i \leq (c^i)^T \bar{x}$ for all $i = 1, 2, \dots$. From conditions (58) and (59), we see that

$$c^T x^* = \lim_{i \rightarrow \infty} (c^i)^T x^i \leq \lim_{i \rightarrow \infty} (c^i)^T \bar{x} = c^T \bar{x}$$

which means that x^* is an optimal solution to problem $P(c; a, b, T)$.

To prove the necessity, let x^* be an optimal solution of problem $P(c; a, b, T)$.

Suppose that $\Omega = \{\omega_1, \omega_2, \dots, \omega_{2n-2}\}$ is an index set such that $\Omega \cap T = \emptyset$. We extend the definition of both $a(t)$ and $b(t)$ to Ω as follows:

$$a(\omega_j) = \begin{cases} (u^j)^T, & \text{for } j = 1, 2, \dots, n-1, \\ -(u^j)^T, & \text{for } j = n, n+1, \dots, 2n-2; \end{cases} \quad (62)$$

and

$$b(\omega_j) = \begin{cases} 1 + (u^j)^T x^*, & \text{for } j = 1, 2, \dots, n-1, \\ 1 - (u^j)^T x^*, & \text{for } j = n, n+1, \dots, 2n-2, \end{cases} \quad (63)$$

where $\{u^1, u^2, \dots, u^{n-1}\}$ is an arbitrarily chosen orthonormal basis for the subspace $\{x \mid c^T x = 0\} \subset \mathbb{R}^n$.

Consider the LSIP problem $P(c; a, b, T \cup \Omega)$ which can be explicitly formulated as follows:

$$\min \quad c^T x \quad (64)$$

$$\text{s. t.} \quad a(t)x \leq b(t), \quad t \in T \quad (65)$$

$$(u^i)^T x \leq 1 + (u^i)^T x^*, \quad i = 1, 2, \dots, n-1 \quad (66)$$

$$-(u^i)^T x \leq 1 - (u^i)^T x^*, \quad i = 1, 2, \dots, n-1 \quad (67)$$

It is obvious that x^* is an optimal solution of problem $P(c; a, b, T \cup \Omega)$. We define, for each $\delta > 0$, a compact set $S_\delta \subset \mathbb{R}^n$ by

$$S_\delta = \{x \in \mathbb{R}^n \mid c^T x \leq c^T x^*, \|x - x^*\| = \delta\}.$$

Consider each $i = 1, 2, \dots$ and the corresponding set $S_{1/(i+1)}$. For each $y \in S_{1/(i+1)}$, we choose an index $t_y^i \in T \cup \Omega$ and an open ball $O(y, \varepsilon_y) \subset \mathbb{R}^n$ with centre y and radius ε_y as follows:

Case I. $y \in S_{1/(i+1)}$ satisfies $c^T y < c^T x^*$. Then, y is infeasible to problem $P(c; a, b, T)$, and y must violate a constraint corresponding to some $t_y^i \in T$. That is, $a(t_y^i)y > b(t_y^i)$. Thus, there exists $O(y, \varepsilon_y)$ such that

$$a(t_y^i)x > b(t_y^i), \quad \text{for all } x \in O(y, \varepsilon_y). \quad (68)$$

Case II. $y \in S_{1/(i+1)}$ satisfies $c^T y = c^T x^*$. In this case, since the point $x^* + \alpha(y - x^*)$ ($\alpha \geq 0$) moves away from x^* as α increases from $\alpha = 0$, it will eventually violate at least one of the constraints in (66) and (67) that correspond to the index set Ω . Suppose that the first of these constraints that is to be violated has index $t_y^i \in \Omega$. Then,

$$a(t_y^i)(y - x^*) = \max \{a(\omega)(y - x^*) \mid \omega \in \Omega, a(\omega)(y - x^*) > 0\}.$$

It is easy to see that the point

$$\bar{y} = x^* + (y - x^*) / (a(t_y^i)(y - x^*)) \quad (69)$$

is on the hyperplane $a(t_y^i)x = b(t_y^i)$. In addition, x^* , y and \bar{y} are on the same line, and

$$\|\bar{y} - x^*\| = \|(y - x^*) / (a(t_y^i)(y - x^*))\| \geq 1 > 1/(i + 1) = \|y - x^*\|.$$

The last inequality, together with (69), means that y is located strictly between x^* and \bar{y} on the line segment joining x^* and \bar{y} . Consider the open ball $O(\bar{y}, 1/(i + 1))$ with centre \bar{y} and radius $1/(i + 1)$. There exists $\varepsilon_y > 0$ such that

$$O(y, \varepsilon_y) \subset \{sx^* + (1 - s)x \mid 0 < s < 1, x \in O(\bar{y}, 1/(i + 1))\}, \quad (70)$$

as the right hand side of (70) is an open set containing y .

Since $S_{1/(i+1)}$ is covered by the family of open sets

$$\{O(y, \varepsilon_y) \mid y \in S_{1/(i+1)}\},$$

it is covered by a finite subfamily

$$\{O(y^j, \varepsilon_{y^j}) \mid j = 1, 2, \dots, n_i\}.$$

This gives a finite sequence

$$\{y^i \mid i = 1, 2, \dots, n_i\} \subset S_{1/(i+1)}.$$

With $t_{y^j}^i$ determined above for $j = 1, 2, \dots, n_i$, let

$$\hat{T}_i = \{t_{y^1}^i, t_{y^2}^i, \dots, t_{y^{n_i}}^i\}.$$

Consider the LP problem $P(c; a, b, \hat{T}_i)$. It is clear that this problem is feasible and bounded. Therefore, according to Theorem 2.2, it has an optimal solution x^i at the vertex of some inclusive region $L(c; a, b, \bar{T}_i)$ of problem $P(c; a, b, \hat{T}_i)$, where

$$\bar{T}_i = \{t_1^i, t_2^i, \dots, t_n^i\} \subset \hat{T}_i. \quad (71)$$

Thus,

$$a(t_j^i)x^i = b(t_j^i), \quad j = 1, 2, \dots, n, \quad (72)$$

and there exist $\lambda_j^i \geq 0$, $j = 1, 2, \dots, n$, such that

$$-c = \sum_{j=1}^n \lambda_j^i a(t_j^i)^T. \quad (73)$$

If $x^i = x^*$ for some $i = i_0$, x^* is the vertex of $L(c; a, b, \bar{T}_{i_0})$. Equations (72) and

(73) become, respectively,

$$a(t_j^{i_0})x^* = b(t_j^{i_0}), \quad j = 1, 2, \dots, n, \quad (74)$$

and

$$-c = \sum_{j=1}^n \lambda_j^{i_0} a(t_j^{i_0})^T, \quad \lambda_j^{i_0} \geq 0, \quad 1 \leq j \leq n. \quad (75)$$

Because none of the constraints corresponding to $\omega \in \Omega$ is active at x^* , (74) implies that $\bar{T}_{i_0} \subset T$. Suppose that there are k ($1 \leq k \leq n$) positive multipliers in (75), and assume without loss of generality that they are $\lambda_j^{i_0}$, $j = 1, 2, \dots, k$. Then (58)–(61) holds true if we take $T_i = \{t_1^{i_0}, t_2^{i_0}, \dots, t_k^{i_0}\}$, $\tilde{x}^i = x^*$, $\tilde{c}^i = c$, and $\lambda_j^i = \lambda_j^{i_0}$ for all $i = 1, 2, \dots$, and $1 \leq j \leq k$.

Now consider $x^i \neq x^*$ for all i . In this case, the half-line starting from x^* and passing through x^i intersects S at some point \bar{x}^i . Since S is covered by

$$\{O(y^j, \varepsilon_{y^j}) \mid j = 1, 2, \dots, n_i\},$$

we see that

$$\bar{x}^i \in \cup \{O(y^j, \varepsilon_{y^j}) \mid j = 1, 2, \dots, n_i\}.$$

If there is some y^j ($1 \leq j \leq n_i$) satisfying $c^T y^j < c^T x^*$ such that $\bar{x}^i \in O(y^j, \varepsilon_{y^j})$, then x^i is located between x^* and \bar{x}^i as \bar{x}^i is infeasible to problem $P(c; a, b, \bar{T}_i)$. This implies that

$$c^T x^i > c^T \bar{x}^i \geq c^T x^* - \|\bar{x}^i - x^*\| = c^T x^* - 1/(i+1).$$

Otherwise, there must be some y^j ($1 \leq j \leq n_i$) satisfying $c^T y^j = c^T x^*$ such that $\bar{x}^i \in O(y^j, \varepsilon_{y^j})$. In this case, x^i is located between x^* and some \hat{y}^i satisfying

$$\hat{y}^i \in O(\bar{y}^j, \frac{1}{i+1}) \cap \{x \mid a(t_{y^j}^i)x = b(t_{y^j}^i)\},$$

where \bar{y}^j is related to y^j according to (69). This implies that

$$c^T x^i \geq c^T \hat{y}^i = c^T (y^j + (\hat{y}^i - y^j)) \geq c^T x^* - \|\hat{y}^i - y^j\| > c^T x^* - 1/(i+1).$$

Thus, x^i satisfies

$$c^T x^* - 1/(i+1) < c^T x^i \leq c^T x^* \quad (76)$$

and the sequence $\{x^i\}$ satisfies

$$\{x^i\} \subset \{x \in \mathbb{R}^n \mid a(\omega)x \leq b(\omega) + 1, \forall \omega \in \Omega, \text{ and } c^T x^* - 1 \leq c^T x \leq c^T x^*\}.$$

This shows that $\{x^i\}$ is bounded and has a converging subsequence which without loss of generality is assumed to be $\{x^i\}$ itself. Let $x^i \rightarrow x^0$. Then from (76) we get $c^T x^0 = c^T x^*$ and x^0 satisfies constraints (66) and (67).

If $x^0 = x^*$, it is easy to see that (58) and (59) of the theorem hold true by taking $\tilde{x}^i = x^i$ and $\tilde{c}^i = c$. (72) indicates that $\bar{T}_i \subset T$ when i is large. Then, (60) and (61) are derived from (72) and (73) by dropping those terms for which $\lambda_j^i = 0$.

If $x^0 \neq x^*$, according to Lemma 4.3, there exist sequences $\{\tilde{x}^i\}$ and $\{\tilde{c}^i\}$ such that (58)–(61) are all satisfied. As $\tilde{x}^i \rightarrow x^*$, when i is sufficiently large, any of the constraints corresponding to $\omega \in \Omega$ is inactive at x^i . Thus, the constraint indices in (60) are contained in T . The proof of Theorem 4.4 is complete. \square

Consider the sequences $\{\tilde{x}^i\}$ and $\{\tilde{c}^i\}$ in Theorem 4.4. Let

$$T(\tilde{x}^i) = \{t \mid t \in T, a(t)\tilde{x}^i \leq b(t)\}.$$

Then, \tilde{x}^i is an optimal solution to problem $P(\tilde{c}^i; a, b, T(\tilde{x}^i))$, and (60) and (61) are equivalent to the complementarity condition for problem $P(\tilde{c}^i; a, b, T(\tilde{x}^i))$ at \tilde{x}^i . It is clear that for any LSIP problem, the classical KKT, complementarity, and Lagrange saddle point conditions either all hold or all fail at an optimal solution. Therefore, we have the following straightforward consequence of Theorem 4.4.

THEOREM 4.5 *Let $x^* \in \mathcal{F}$. The following statements are equivalent:*

- (i) $x^* \in \mathcal{F}^*$;
- (ii) (generalized KKT condition) there exist sequences $\{\tilde{c}^i\}$ and $\{\tilde{x}^i\}$ such that

$$\tilde{c}^i \rightarrow c, \quad (i \rightarrow \infty), \quad (77)$$

$$\tilde{x}^i \rightarrow x^*, \quad (i \rightarrow \infty), \quad (78)$$

and

$$-\tilde{c}^i \in \mathcal{A}(\tilde{x}^i), \quad i = 1, 2, \dots; \quad (79)$$

(Note that $\mathcal{A}(\tilde{x}^i)$ as defined in Theorem 4.1 is the same as the active cone at \tilde{x}^i w.r.t. problem $P(\tilde{c}^i; a, b, T(\tilde{x}^i))$.)

- (iii) (generalized complementarity condition) there exist sequences $\{\tilde{c}^i\}$ and $\{\tilde{x}^i\}$ satisfying (77) and (78), and $t_j^i \in T(\tilde{x}^i)$, $1 \leq j \leq k$, such that $\{a(t_j^i) \mid 1 \leq j \leq k\}$ is linearly independent,

$$a(t_j^i)\tilde{x}^i = b(t_j^i), \quad 1 \leq j \leq k, \quad i = 1, 2, \dots$$

and

$$-\tilde{c}^i = \sum_{j=1}^n \lambda_j^i a(t_j^i)^T, \quad \lambda_j^i > 0, \quad j = 1, 2, \dots, k, \quad i = 1, 2, \dots;$$

- (iv) (generalized Lagrange saddle point condition) there exist sequences $\{\tilde{c}^i\}$ and $\{\tilde{x}^i\}$ satisfying (77) and (78), and $\lambda_i \in \mathbb{R}_+^{(T(\tilde{x}^i))}$, $i = 1, 2, \dots$, such that

$$L(\tilde{x}^i, \lambda) \leq L(x^i, \lambda_i) \leq L(x, \lambda_i), \quad i = 1, 2, \dots$$

for all $x \in \mathcal{F}_i = \{x \mid a(t)x \leq b(t) \text{ for all } t \in T(\tilde{x}^i)\}$ and all $\lambda \in \mathbb{R}_+^{(T(\tilde{x}^i))}$,

where

$$L(x, \lambda) = c^T x + \sum_{t \in T} \lambda(t)(a(t)x - b(t)).$$

□

As an application, a proof of Theorem 4.2 based on Theorem 4.5 is given below.

Proof of Theorem 4.2: Let $x^* \in \mathcal{F}^*$. Let \tilde{c}^i , \tilde{x}^i , t_j^i and λ_j^i be as in Theorem 4.5(iii). Take

$$\varepsilon_i = \max \left\{ c^T x^* - (\tilde{c}^i)^T \tilde{x}^i, 0 \right\}$$

and $\lambda_i \in R_+^{(T)}$ such that

$$\lambda_i(t) = \begin{cases} \lambda_j^i, & t = t_j^i, j = 1, \dots, k, \\ 0, & \text{otherwise.} \end{cases}$$

By the assumptions, $\varepsilon_i \rightarrow 0_+$ and

$$\sum_{t \in T} \lambda_i(t) a(t) = -\tilde{c}^i \rightarrow -c,$$

which justifies (15). Moreover,

$$\begin{aligned} \sum_{t \in T} \lambda_i(t) b(t) &= \sum_{j=1}^k \lambda_j^i b(t_j^i) \\ &= \sum_{j=1}^k \lambda_j^i a(t_j^i) \tilde{x}^i \\ &= -(\tilde{c}^i)^T \tilde{x}^i \\ &\leq \varepsilon_i - c^T x^*, \end{aligned}$$

so that (14) holds.

Now we assume that (14) and (15) hold. Let $\tilde{c}^i := -\sum_{t \in T} \lambda_i(t) a(t)$ and let $x \in \mathcal{F}$. Then,

$$(\tilde{c}^i)^T x = -\sum_{t \in T} \lambda_i(t) a(t) x \geq -\sum_{t \in T} \lambda_i(t) b(t) \geq c^T x^* - \varepsilon_i \geq c^T x^*.$$

Taking limits as $i \rightarrow \infty$ we get that $c^T x \geq c^T x^*$. Thus, $x^* \in \mathcal{F}^*$. □

Now we go back and revisit Examples 1 and 2. As we have seen earlier in this section, the classical optimality conditions of Theorem 4.1 are not satisfied at the optimal solutions for both examples. On the other hand, the LSIP problems in both examples satisfy the optimality conditions given in Theorem 4.5.

We can easily verify that, for Example 1, Theorem 4.4 is satisfied with the fol-

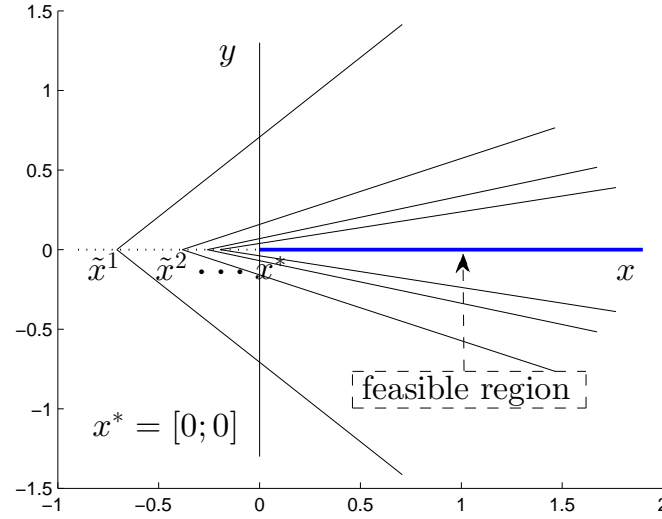


Figure 2. Explanation of Theorem 4.5 by Example 1.

lowing sequences:

$$\begin{aligned}\tilde{x}^i &= [-\sin(\pi/4i); 0], \\ \tilde{c}^i &= c = [1; 0], \\ T_i &= \{\pi/4i, (4i-1)\pi/4i\}, \quad (k=2, \quad t_1^i = \pi/4i, \quad t_2^i = (4i-1)\pi/4i).\end{aligned}$$

Note that in the above example, $\{\tilde{c}^i\}$ can be chosen as a constant sequence, but $\{\tilde{x}^i\}$ cannot be constant. Figure 2 shows the above sequence $\{\tilde{x}^i\}$ and the active constraints at \tilde{x}^i for the first few i values. For this example, there are many other choices for the sequences $\{\tilde{x}^i\}$, $\{\tilde{c}^i\}$, and $\{T_i\}$.

For Example 2, any point on the negative x -axis is an optimal solution. For each optimal solution $x^* = [-\alpha; 0]$, (58)–(61) of Theorem 4.4 are satisfied with the following \tilde{x}^i , \tilde{c}^i and T_i (see Figure 3):

$$\begin{aligned}\tilde{x}^i &= [-\alpha; -\alpha \tan(\pi/5i)], \quad i = 1, 2, \dots, \\ \tilde{c}^i &= c = [\sin(\pi/5i); -\cos(\pi/5i)], \quad i = 1, 2, \dots, \text{ and} \\ T_i &= \{\pi/5i\}, \quad (k=1, \quad t_1^i = \pi/5i), \quad i = 1, 2, \dots.\end{aligned}$$

It is of interest to note that optimal solutions of LSIP problems can be classified into four types according to the convergence pattern of the sequences $\{\tilde{x}^i\}$ and $\{\tilde{c}^i\}$, as specified below.

- I. $x^* \in \mathcal{F}^*$ and (ii)-(iv) of Theorem 4.5 are satisfied by $\tilde{x}^i = x^*$ and $\tilde{c}^i = c$ for $i = 1, 2, \dots$.
- II. $x^* \in \mathcal{F}^*$ is not of type I, (ii)-(iv) of Theorem 4.5 are satisfied by $\tilde{c}^i = c$, $i = 1, 2, \dots$, and some non-constant sequence $\{\tilde{x}^i\}$.
- III. $x^* \in \mathcal{F}^*$ is not of type I or type II, (ii)-(iv) of Theorem 4.5 are satisfied by $\tilde{x}^i = x^*$, $i = 1, 2, \dots$, and some non-constant sequence $\{\tilde{c}^i\}$.
- IV. $x^* \in \mathcal{F}^*$ is not any of above type. In this case, none of the sequences $\{\tilde{x}^i\}$ and $\{\tilde{c}^i\}$ in Theorem 4.5 can be constant.

Type I optimal solutions are those of LSIP problems at which the classical KKT, complementarity and Lagrangian saddle point conditions hold. An example of type II optimal solution is the optimal solution of the LSIP problem in Example 2. The optimal solution $x^* = [0; 0]$ for the problem in Example 1 is a type III optimal solution, while the remaining optimal solutions for this problem are all of type IV.

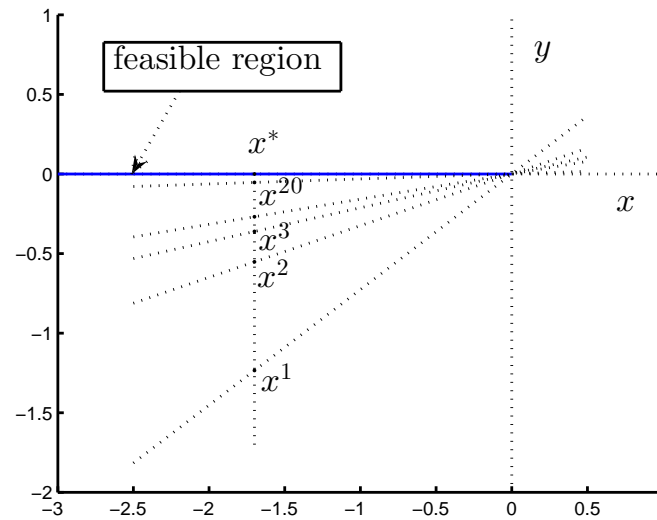


Figure 3. Explanation of Theorem 4.5 by Example 2.

5. Comments

In this paper, three new equivalent conditions characterizing the optimality of a feasible solution for general linear semi-infinite programming without constraint qualifications have been provided. They are direct generalizations of the classical KKT, complementarity, and Lagrangian saddle point conditions. The method used in this paper, in contrast to the classical convex analysis method, is based on finite dimensional analysis and the geometric properties of so called inclusive cone that coexists with the associated inclusive region. The current paper and some related previous works ([10–12]) demonstrate that the inclusive cone and inclusive region are efficient tools for the theory and computation of linear optimization problem. The method is expected to be useful in the theoretical analysis and the numerical study of general convex optimization. This will be investigated as a future research.

References

- [1] [Dinh N, Goberna MA, López MA, Volle M. Convex inequalities without constraint qualification nor closedness condition, and their applications in optimization. Set-valued Var. Anal. 2010; 18: 423–445.](#)
- [2] Goberna MA. Linear semi-infinite optimization: a guided tour. IMCA Monographs No. 40. IMCA, Lima (Perú); 2005.
- [3] [Goberna MA, Lopez MA. A comprehensive survey of linear semi-infinite optimization. In: Reemtsen R, Ruckmann JJ, editors. Semi-Infinite Programming: Nonconvex Optim Appl; vol. 25; Dordrecht: Kluwer; 1998. p. 3–27.](#)

- [4] Goberna MA, Lopez MA. [Linear semi-infinite programming theory: An updated survey. European Journal of Operational Research. 2002; 143: p. 390–405.](#)
- [5] Goberna MA, Lopez MA, and Todorov MI. [Extended active constraints in linear optimization with applications, SIAM J. Optim. 2003; 14: p. 806–619.](#)
- [6] Hettich R, Kortanek KO. [Semi-infinite programming: Theory, method, and applications. SIAM Review, 1993; 35: p. 380–429.](#)
- [7] Klatte D, Li W. [Asymptotic constraint qualifications and global error bounds for convex inequalities. Mathematical Programming, Ser. B. 1999; 84: p. 137–160.](#)
- [8] Kortanek KO, Strojwas HM. [On constraint sets of infinite linear systems over ordered fields. Mathematical Programming, 1985; 33: p. 146–161.](#)
- [9] Kortanek KO, Zhang Q. [Perfect duality in semi-infinite and semi-definite programming. Mathematical Programming. 2001; 91: p. 127–144.](#)
- [10] Liu Y. [An exterior point method for linear programming based on inclusive normal cones. JOURNAL Industrial and Management Optimization. 2010; 6: p. 825–846.](#)
- [11] Liu Y. [Duality in linear programming: from trichotomy to quadrichotomy. Journal of Industrial and Management Optimization. 2011; 7: p. 1003–1011.](#)
- [12] Liu Y, Ding MF. [A ladder method for semi-infinite programming. Journal of Industrial and Management Optimization. 2014; 10: P. 397–412.](#)
- [13] Lopez MA, Still G. [Semi-infinite programming. European Journal of Operational Research. 2007; 143: p. 491–518.](#)
- [14] Lopez MA, Vercher E. [Optimality conditions for non-differentiable convex semi-infinte programming Mathematical Programming. 1983; 27: p. 307–319.](#)
- [15] Puente R, Vera De Serio V. [Locally Farkas-Minkowski Linear Semi-Infinte Systems. TOP. 1999; 7: p. 103–121.](#)
- [16] Rubio JE. [Optimal control problems with unbounded constraint sets. Optimization. 2000; 48: p. 191-210.](#)
- [17] Shapiro A. [Semi-infinite programming, duality, discretization and optimality conditions. Optimization. 2009; 58: p. 133–161.](#)
- [18] Slupphaug O, Imsland L, Foss A. [Uncertainty modelling and robust output feedback control of nonlinear discrete systems: A mathematical programming approach. Int. J. Robust Nonlinear Control. 2000; 10: p. 1129–1152.](#)